Survey of commercial technologies for face recognition in Video

Prepared by:
Dmitry O. Gorodnichy, Eric Granger, Paolo Radtke
Canada Border Services Agency
Ottawa ON
Canada K1A 0L8

Contract Scientific Authority: Pierre Meunier
DRDC Centre for Security Science
613-992-0753

IMPORTANT INFORMATIVE STATEMENTS

PROVE-IT (FRiV) Pilot and Research on Operational Video-based Evaluation of Infrastructure and Technology: Face Recognition in Video project, PSTP 03-401BIOM, was supported by the Canadian Safety and Security Program (CSSP) which is led by Defence Research and Development Canada's Centre for Security Science, in partnership with Public Safety Canada. Led by Canada Border Services Agency partners included: Royal Canadian Mounted Police, Defence Research Development Canada, Canadian Air Transport Security Authority, Transport Canada, Privy Council Office; US Federal Bureau of Investigation, National Institute of Standards and Technology, UK Home Office; University of Ottawa, Université Québec (ÉTS).

The CSSP is a federally-funded program to strengthen Canada's ability to anticipate, prevent/mitigate, prepare for, respond to, and recover from natural disasters, serious accidents, crime and terrorism through the convergence of science and technology with policy, operations and intelligence.

Science and Engineering Directorate

Border Technology Division

Division Report 2014-22 (TR)
June 2014

Survey of commercial technologies for Face Recognition in Video

D. Gorodnichy, E. Granger, P. Radtke

This page left intentionally blank

Abstract

This report reviews commercial off-the-shelf (COTS) solutions and related patents for face recognition in video surveillance applications. Commercial products are analyzed using such criteria as processing speed, feature selection techniques, ability to perform screening against the watch list, and ability to perform both still-to-video and video-to-video recognition.

Keywords: video-surveillance, face recognition in video, instant face recognition, watch-list screening, biometrics, reliability, performance evaluation

Community of Practice: Biometrics and Identity Management

Canada Safety and Security (CSSP) investment priorities:

1. Capability area: P1.6 – Border and critical infrastructure perimeter screening technologies/ protocols for rapidly detecting and identifying threats.
2. Specific Objectives: O1 – Enhance efficient and comprehensive screening of people and cargo (identify threats as early as possible) so as to improve the free flow of legitimate goods and travellers across borders, and to align/coordinate security systems for goods, cargo and baggage;
3. Cross-Cutting Objectives CO1 – Engage in rapid assessment, transition and deployment of innovative technologies for public safety and security practitioners to achieve specific objectives;
4. Threats/Hazards F – Major trans-border criminal activity – e.g. smuggling people/ material

Acknowledgements

Disclaimer

In no way do the results presented in this paper imply recommendation or endorsement by the Canada Border Services Agency, nor do they imply that the products and equipment identified are necessarily the best available for the purpose. The information presented in this report contains only the information available in the public domain.

Release Notes

Context: This document is part of the set of reports produced for the PROVE-IT(FRiV) project. All PROVE-IT(FRiV) project reports are listed below.

- Dmitry Gorodnichy, Eric Granger "PROVE-IT(FRiV): framework and results". Also published in Proceedings of NIST International Biometrics Performance Conference (IBPC 2014), Gaithersburg, MD, April 1-4, 2014. Online at http://www.nist.gov/itl/iad/ig/ibpc2014.cfm.
- Dmitry Gorodnichy, Eric Granger, "Evaluation of Face Recognition for Video Surveillance". Also published in Proceedings of NIST International Biometric Performance Conference (IBPC 2012), Gaithersburg, March 5-9, 2012. Online at http://www.nist.gov/itl/iad/ig/ibpc2012.cfm.
- E. Granger, P.Radtke, and D. Gorodnichy, "Survey of academic research and prototypes for face recognition in video,"
- D. Gorodnichy, E.Granger, and P.Radtke, "Survey of commercial technologies for face recognition in video,"
- E. Granger and D. Gorodnichy, "Evaluation methodology for face recognition technology in video surveillance applications,'
- D. Gorodnichy, E. Granger, E. Choy, W. Khreich, P.Radtke, J. Bergeron, and D. Bissessar, "Results from evaluation of three commercial off-the-shelf face recognition systems on Chokepoint dataset,"
- S. Matwin, D. Gorodnichy, and E. Granger, "Using smooth ROC method for evaluation and decision making in biometric systems,"
- E. Neves, S. Matwin, D. Gorodnichy, and E. Granger, "3D face generation tool Candide for better face matching in surveillance video,"
- E. Neves, S. Matwin, D. Gorodnichy, and E. Granger, "Evaluation of different features for face recognition in video,"

The PROVE-IT(FRiV) project took place from August 2011 till March 2013. This document was drafted and discussed with project partners in March 2013 at the Video Technology for National Security (VT4NS) forum. The final version of it was produced in March 2014.

Typesetting: All tabulated content in this report was produced automatically using LATEX content for improved source control, flexibility and maintainability. The report contains automatically generated hyper-link references and table of contents for easier navigation and reading on-line. All references having an "http" prefix provide virtual links to the corresponding internet pages.

Contact: Correspondence regarding this report should be directed to DMITRY dot GORODNICHY at CBSA dot GC dot CA.

Contents

List of Figures

List of Tables

1 Introduction

As discussed in report [1] and illustrated in Figure 1, developing face recognition (FR) solutions for video surveillance applications requires implementation and integration of many *Face Processing* tasks [1]. While the most critical Face Processing module is the *Face Matching* module, it is the integration of all modules that makes a FR solution successful for a given application.

There are over a hundred companies referenced on the internet, some of which are listed in Table 1, that provide face recognition solutions. Of these, there are only a few, referred to as *FR developers* in the table, that provide their own FR matching components. As indicated in the report [1], three main groups of FR matching products are recognized:

1. Technology developed for high-performance in still-to-still comparison, such as NEC, Morpho (with its acquired Sagem and L1 solutions), Cognitec.

2. Technology developed for high-performance in low resolution or video-to-video comparison, such as Google-acquired PittPatt, and

3. Technology that is less performing but which offers more affordable and easier to integrate options, such as Neurotechnology.

Figure 2 provides a comparative performance analysis from [2] of three FR matching products listed above, represented in terms of the Detection-Error Tradeoff (DET) curves that plot False Non-Match Rate (FNMR) against False Match Rate (FMR): NEC (light blue curves - best performing), Cognitec (dark blue curves - second best), and PittPatt (red lines).

The majority of other companies are *FR integrators*, who either license FR matching technology from FR developers or build their solutions using the Open Source libraries, of which there are several available on the internet. A non-exhaustive list of FR developers, FR integrators and Open Sources FR libraries is provided in Table 1.

In this report, we present an overview of these FR products (Section 2) and related patents (Section 3), in the context of their applicability for video surveillance applications, and provide recommendations based thereon on the selection of COTS products for further testing and piloting (Section 4).

The methodology for testing COTS FR products is developed in [3]. The results from testing several COTS FR products are presented in [4]. Finally, the survey of academic solutions, which provides more detail on the FR approaches mentioned below is presented in [1].

---

[1]For the definitions and analysis of face processing tasks, see "Introduction to the First IEEE Workshop on Face Processing in Video" at http://www.visioninterface.net/fpiv04/preface.html.

Table 1: Face recognition developers, integrators, and open source libraries

FR developers	Website
Acsys Biometrics	www.acsysbiometrics.com.
Animetrics	www.animetrics.com
Ayonix	www.ayonix.com
Bayometric	www.bayometric.com
Behrooz Kamgar-Parsi	www.biometrics.org/bc2006/presentations/Tues_Sep_19/BSYM/19_Kamgas-Parsi_research.pdf
Betaface	www.betaface.com
Cognitec Systems GmbH	www.cognitec-systems.de
Cross Match Technologies, Inc.	www.crossmatch.com
Cybula Ltd.	www.cybula.com
Face.com	www.face.com
Facial Forensic (F2)	www.faceforensics.com
L-1 Identity Solutions, Inc. (acquired Viisage and Identix)	www.l1id.com
Luxand, Inc.	www.luxand.com
Morpho (acquired L1, 2011)	www.morpho.com
NeoFace - NEC	www.necam.com/Biometrics/doc.cfm?t=FaceRecognition
Neuro Technology	www.neurotechnology.com
OmniPerception	www.omniperception.com
PittPatt: Pittsburgh Pattern Recognition (acquired by Google)	www.pittpatt.com
Sensible Vision, Inc.	www.sensiblevision.com
FR integrators	
Advanced Corp. Security Systems	www.acss.co.za
Airborne Biometrics Group	www.facefirst.com
Arti-Vision	www.arti-vision.com
Aurora	www.facerec.com
Avalon Biometrics	www.avalonbiometrics.com
Canadian Bank Note	www.cbnco.com
Csystems Advanced Biometrics	www.ex-sight.com
EAL	www.eal.nl
Face.com developers	www.developers.face.com
Facing-IT	www.facing-it.com
Guardia	www.guardia.dk
Herta Security	www.hertasecurity.com
ID One, Inc.	www.idoneinc.com
IITS, S.L.	www.iits.se
INO	www.ino.ca
Intelligent Security Systems	www.isscctv.com
IntelligenTek	www.intelligentek.com
Inttelix	www.inttelix.com
iView	www.iviewsystems.com
iWT	www.iwtek.net
Kee Square	www.keesquare.com
Kiwi Security	www.kiwi-security.com
NextgenId	www.nextgenid.com/
NICTA	www.nicta.com.au/
Omron	www.omron.com
Panvista	www.panvista.com
PSP Security	www.pspsecurity.com
Quantum Signal	www.quantumsignal.com
TAB Systems	www.tab-systems.com
The Covenant Consortium (TCC)	www.tcc.us.com
XID Technologies Pte Ltd.	www.xidtech.com
IntelliVision	www.intelli-vision.com
Open Source FR codes	
CSU: Evaluation of Face Recognition	www.cs.colostate.edu/evalfacerec
CSU: FaceL: Facile Face Labeling	www.cs.colostate.edu/facel
CSU: Baseline 2010 Algorithms	www.cs.colostate.edu/facerec/algorithms/baselines2011.php
RTFTR: Real-Time Face Tracking and Recognition	rtftr.sourceforge.net/
Face Recognition using Associative Neural Networks	www.videorecognition.com/FRiV
OpenCV Face Recognition	docs.opencv.org/modules/contrib/doc/facerec/facerec_tutorial.html
Candide 3D model-based coding of human faces	www.icg.isy.liu.se/candide/
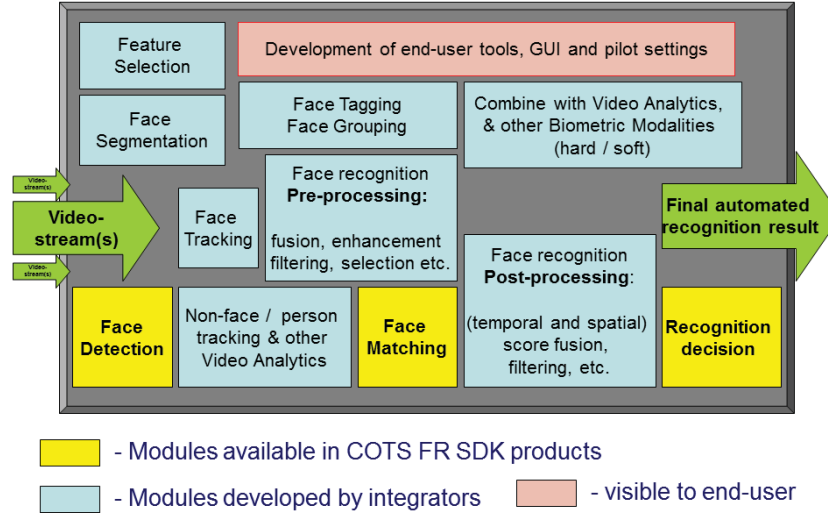
Figure 1: Face processing tasks required for integration of FR solutions into a video surveillance application.

2   Commercial products for face recognition in video

A non-exhaustive list of FR solution providers is available at the Biosecure website[2], which was our primary source when analyzing commercial products.

Since the focus of this study is on COTS and patents that are applicable to video surveillance, the technologies and patents that do not deal with video-based applications, such as those dealing with face recognition for access control applications only, were not investigated. Even though some solution providers do provide numbers for the matching performance, this performance, unless explicitly mentioned, is obtained using standard still image face datasets, rather than video-based datasets.

Table 2 summarizes COTS FR products that were found relevant to the subject of this study, specifically to the seven video surveillance applications studied in PROVE-IT(FRiV) project listed below:

1. screening of faces (screening against wanted list);
2. fusion of face recognition from different cameras;
3. face recognition-assisted tracking;
4. matching a face/person across several video feeds;
5. multi-modal recognition (e.g. face and voice)
6. soft-biometric based tracking/recognition

The following requirements were used to analyze commercial products:

---

[2]http://biosecure.it-sudparis.eu/AB/index.php?option=com_content&view=article&id=21&Itemid=26
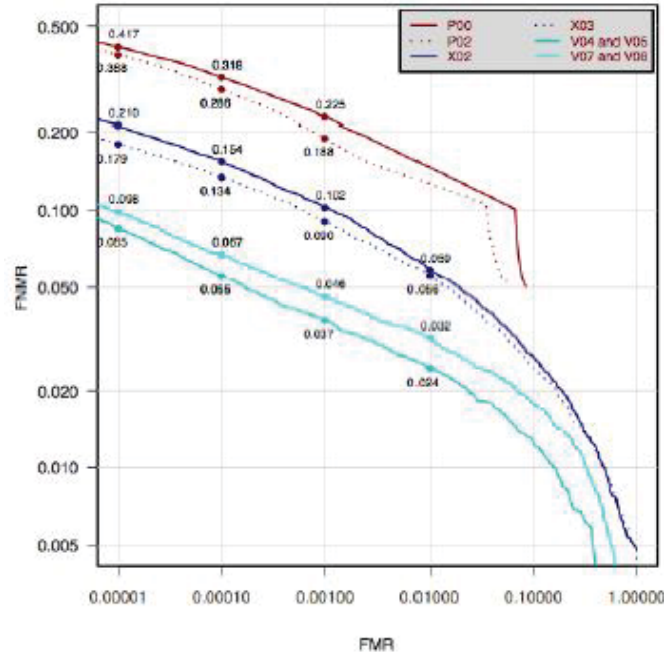
Figure 2: The Detection Error Tradeoff (DET) curves for NEC, Cognitec and PittPatt products (from NIST Multiple-Biometric Evaluation).

- **Processing speed**: a real-time face recognition in video system should be able to run at 30 fps over several cameras.
- **Restrained watch list**: to reduce memory usage and accelerate matching speed the system must support a restrained watch list.
- **Feature extraction/selection**: the used techniques must be robust for unconstrained environments, which are subject to changes in pose, lighting, capture from diverse video equipment.
- **Functions**: the system must be able to perform both still-to-video and video-to-video recognition to be applicable for the video surveillance applications listed above.

Table 2 also presents information related to the memory usage for templates, matching speed (faces per second) and head orientation.

11

Table 2: COTS software solutions for face recognition in video (* requires extra coding, $^{+}$ requires complementary SDKs, × available by special request).

Name	Company	Approach	Matching	Set	Tracking	Application	Template size	Match speed	Head yaw (degrees)
NeoFace Suite SDK	NEC http://nec.com	both	template	both	multiple	1, 4, 5^{+}	–	1000000/s	+30 to -30
VeriLook Surveillance SDK	NEUROtechnology http://neurotechnology.com/	both	template	both	multiple	1, 2, 3, 4, 5^{+}	4kb – 36kb	60000/s	+45 to -45
FaceR	Animetrics http://animetrics.com/	still-to-still	template	both	no	1, 2 3*	6kb	–	+45 to -45
FaceIT SDK	L1 http://www.l1id.com	still-to-still	template	both	no	2, 4^{+}	648b – 7kb	60000000/s	–
PittPatt SDK	Google http://pittpatt.com/	both	classifier	both	multiple	1, 2, 3	10kb – 120kb	–	+18 to -18 +36 to -36
FaceVACS SDK	Cognitec Systems http://cognitec-systems.de	both	template	both	multiple	1, 2, 3, 5^{+}	1424b – 9505b	142000/s (single thread)	+15 to -15
SureMatch 3D Suite for Watchlist	Genex (Technest) http://www.genextech.com	still-to-still	template	open	no	2	–	–	+90 to -90
Notiface II	FACE-TEK http://www.face-tek.com	still-to-still	undisclosed	open	no	2	–	–	–
Acsys FRS SDK	Acsys Biometrics http://www.acsysbiometrics.com	video-to-video	template	open	multiple	1, 2 3	4kb and 8kb	100000/s (still) 25000/s (video)	+15 to -15
FaceFirst	Airbone Biometrics Group, Inc. http://www.facefirst.com	still-to-video	template	open	no	2	1424b – 9505b	142000/s	+15 to -15

2.1 NEC NeoFace Suite

- Vendor: NEC
- Web: http://nec.com
- Type: SDK
- Tasks: facial image matching, CCTV watchlist screening, and searching archived video.

NEC **NeoFace** suite [3] offers face recognition solutions to support and optimise surveillance, identification, and security operations such as monitoring the movement and volume of people in public areas. **NeoFace** showed the best performance the still face recognition problem in the Multiple Biometric Grand Challenge (MBGC) held by NIST in 2008-2009 (see Figure 2). According to NEC, **NeoFace** provides the fastest matching capability (up to 1 million faces/sec) and face matching is still possible with dark glasses, badly lit areas, shadowing, varying angles and partial face cover.

The detection module is based on Generalized Learning Vector Quantization (GLVQ) and Facial Shape Model, while recognition is based on neural network technology [4]. For best results NEC recommends more than 100 pixels between eyes, with less than 15 vertical degrees and less than 30 horizontal degrees.

NeoFace suite comprises several components including **NeoFace Watch**, which allows one to extract and match against a watchlist of individuals, and can be integrated with existing surveillance systems. **NeoFace Match** is another component designed to match photographs against large digital databases of facial images by ranking the database images against the probe image. **NeoFace Find** is designed to search for specific individuals in large volumes of video footage.

NeoFace suite provides an SDK with a runtime license for developer and user environments. The developer SDK (with one development license on a single PC) includes the face recognition library, sample programs, accuracy evaluation tool, and manuals. It allows the development of watchlist-based video surveillance applications under both Windows (Microsoft Visual C++) and Linux (g++ (GCC) 4.1 to 4.7). On the other hand, the user SDK includes a detection and matching license.

2.2 Cognitec FaceVACS SDK

- Vendor: Cognitec Systems
- Web: http://cognitec-systems.de
- Type: SDK and application for integration
- Tasks: face annotation, face identification, enrolment from video to track an individual.

Manufactured by *Cognitec Systems*, the **FaceVACS SDK** is the basis of a family of off-the-shelf products. The most relevant product is **FaceVACS-Video Scan**, a watc- list screening application that provides

---

[3]nz.nec.com/en_NZ/pdfs/NEC_Biometrics_NeoFace.pdf

[4]Yusuke Morishita and Hitoshi Imaoka.'Facial Feature Detection using Generalized LVQ and Facial Shape Model'. MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN.

a basic API for system integrators under the BioAPI[5] specification. The API itself provides all functions available in the **Video Scan** application, such as face tracking, face identification/recognition and enrolment. Camera support ranges from simple USB web-cams to specialized IP cameras. The vendor makes no claim for fusion of many streams, so it can be assumed that each camera is processed independently. The SDK supports both C and C++, with full .Net support for Windows, which virtually allows the development on any language that produces .Net bytecode. One drawback is that **FaceVACS** requires a full DBMS system, supporting *Oracle 11g*, *IBM DB2* and *Microsoft SQL Server*.

Enrolment is done from still images or video sequences, supporting full video-to-video applications. There are no references for the feature type used by the **FaceVACS SDK**, but template size ranges from 1424 to 9505 bytes. The system can match up to 142000 templates per second (single thread), but face rotation (yaw) is limited between +15 to -15 degrees.

The vendor provides some performance measures for both the FERET dataset (closed set problem) and on two open-set scenarios, reaching 95% and 98% true positive rate at 10% false positive rate. However, one scenario is described as a passport issuance point, and the second scenario is described as a typical access control point. We can assume that both scenarios are controlled and provide good, if not ideal, lighting conditions and those results do not reflect a realistic unconstrained video surveillance scenario.

2.3 L1 FaceIT SDK

- Vendor: L1 Identity Solutions [6]
- Web: http://www.l1id.com
- Type: SDK and related solutions
- Tasks: face identification, FRiV for multi-modal recognition.

The **FaceIT SDK** for face recognition is part of a large family of biometric products developed by *L1 Identity Systems*, which also provides solutions for fingerprint/palm, through the **TouchPrint Live Scan Advanced SDK**, and iris recognition, through the **SIRIS SDK**, to provide multi-biometric recognition, which is also supported by the **ABIS System**, a server-side solution to manage biometric data. Besides SDKs for third party developers, the company also provides several off-the-shelf products for integration. All SDKs are available primarily for Windows for C/C++ development (supporting Visual Studio platform), but a Linux version is also available on special request.

The **FaceIT SDK** supports camera captures, still images and individual frames from AVI/MPEG video files. The product description says that it combines facial geometry and skin texture for higher accuracy, which suggests that it uses physiological local features for graph matching. A user template can take from

---

[5]http://www.bioapi.org/

[6]L-1 Identity Solutions, Inc. is a large American defense contractor in Connecticut. It was formed on August 29, 2006, from a merger of Viisage Technology, Inc. and Identix Incorporated. It specializes in selling face recognition systems, electronic passports and other biometric technology to governments such as the United States and Saudi Arabia. It also licenses technology to other companies internationally, including China. On July 26, 2011, Safran (Sagem, Morpho) acquired L-1 Identity Solutions, Inc. for a total cash amount of USD 1.09 billion (Source: Wikipedia).

648 bytes to 7 kbytes, and the system supports 1-to-many identification applications. Finding a head in a frame takes at most 300 ms and 1-to-many template matching speed with the 7kbytes template is of 10 million templates in one second. Typical applications demonstrated on the L1 website indicate that the SDK is suitable for security checkpoint verification and identification. Using the SDK for covert watch-list based screening is unlikely, as the product is tailored to identify/recognize one user at a time from a still image/frame.

The vendor does not provide performance analysis, only stating that it was the best all-around performer on the *Facial Recognition The vendor Test* (FRVT) 2006 sponsored by the *National Institute of Standards and Technology* (NIST) [7]. SDKs are licensed for development and run-time licenses are required for deployment. Known L1 clients include the US Department of State, the Mexican government, the Swedish National Police and the New York Police Department, among many others.

2.4 Neurotechnology VeriLook Surveillance SDK

- Vendor: NEUROtechnology
- Web: http://neurotechnology.com
- Type: SDK
- Tasks: face annotation, face identification and enrolment from video to track an individual.

The **VeriLook Surveillance SDK** allows the development of watchlist-based video surveillance applications in Windows (C++, C# and Visual Basic.Net) and GNU Linux (C++). The SDK is based on the **VeriLook SDK**, which targets the development of face recognition (1:1, closed set) and face identification (1-to-many, open set) applications. Both are part of a family of SDKs targeting biometric application development, including **VeriFinger SDK** for fingerprints, **VeriEye SDK** for iris, **VeriSpeak SDK** for voice recognition and the **MegaMatcher SDK**, which targets the development of large-scale multi-modal biometric applications on a client-server architecture. Both a free demonstration and 30-day SDK trials are available for download at *http://neurotechnology.com*. Besides desktop-based applications, the **VeriLook Embedded SDK** allows the development of face recognition applications on mobile devices running Android 2.2 or higher.

The VeriLook Surveillance SDK supports enrolment from image files, real-time video sequences or previously captured video files. Each user in the system is modelled as a separated template, allowing the fast incremental update of the user template database (the vendor claims less than one second for still images). User template size may be adjusted to provide faster recognition (small template size of 4 kb) or higher accuracy (large template size of 36kb). User template features are undisclosed by the vendor, but the template matching suggests local features. Templates are stored on a SQLite database, which is both fast, cross-platform and portable, requiring no complementary full DBMS installation. Probe matching speed is of 0.5 seconds for a gallery of 30000 templates using a Core i7-2600 processor. Tracking 5 faces on video with a Core i7-2600 processor results on 14fps in real time, below the expected 30 fps. Camera resolution

---

[7]http://www.nist.gov/itl/iad/ig/frvt-home.cfm

plays an important role as well, as the vendor indicates a minimum required distance of 40 pixels between eyes for best recognition. This parameter can be changed to accept smaller distances, but the vendor does not recommend such scenario. Head rotation (yaw) is limited between +45 and -45 degrees, as the SDK requires that both eyes are visible.

The VeriLook Surveillance SDK accuracy is not discussed by the vendor, but the VeriLook SDK brochure details still face recognition performance on the NIST's *Face Recognition Grand Challenge* data set (see Figure 2). Two experiments were performed, one with only one image per subject, and another with four images per subject. At 0.1% false positive rate, the VeriLook SDK false negative rate on three different data sets ranged between 0.92% to 2.46% with one training image per subject, and between 0.04% and 0.06% when using four training images per subject. The vendor claims that the platform is robust to full occlusion, but further testing with the demonstration application shows that the tracking algorithm can handle only some types of full face occlusion. The platform failed with two moving targets that crossed in the camera's field of view, which is very common in crowded scenes. The occluded target was lost and reacquired later as if it were a newly detected individual.

Licensing costs are divided in two categories. The SDK itself costs 790 euro and may be used with no restrictions by all on-site developers. Deployment of products requires another license, which is purchased per computer, with prices varying according to the volume from 290 euro (1 to 10 computers) to 79 euro (4000 to 7999 computers). For larger volumes the vendor needs to be contacted directly.

2.5 Animetrics FaceR

- Vendor: Animetrics
- Web: http://animetrics.com
- Type: Off-the-shelf applications and SDK
- Tasks: face annotation, face identification, enrolment from video to track an individual and FRiV for multi-modal recognition.

Animetrics provides the **FaceR** family of products, based on a proprietary technology that converts 2D pictures to 3D models. The client list is undisclosed, but news clippings suggest the US Department of Homeland Security, US Army and several police departments in different countries use this technology. The base product is the **FaceR Identity Management System** (FIMS), a web server-based facial biometric identity management system that supports single or many thin clients via host-based or cloud-based computing. This approach allows the use of very simple devices as clients, like tablets or smart-phones. FIMS is available both on MySQL or NoSQL (MongoDB), and its architecture is detailed in Fig. 3. FIMS provides a service over a network that can process images provided by many devices and perform face recognition on them. It supports incremental user enrolling and licensing options are based on the number of required user enrolments, starting from a 10,000 users enrolment license.

Animetrics provides a set of complementary applications to FIMS, making **VideoID** very suitable for video surveillance applications. This web-based application allows one to use web cams and IP cameras to stream a video sequence and display resulting face matches and analysis. **VideoID** can be used either
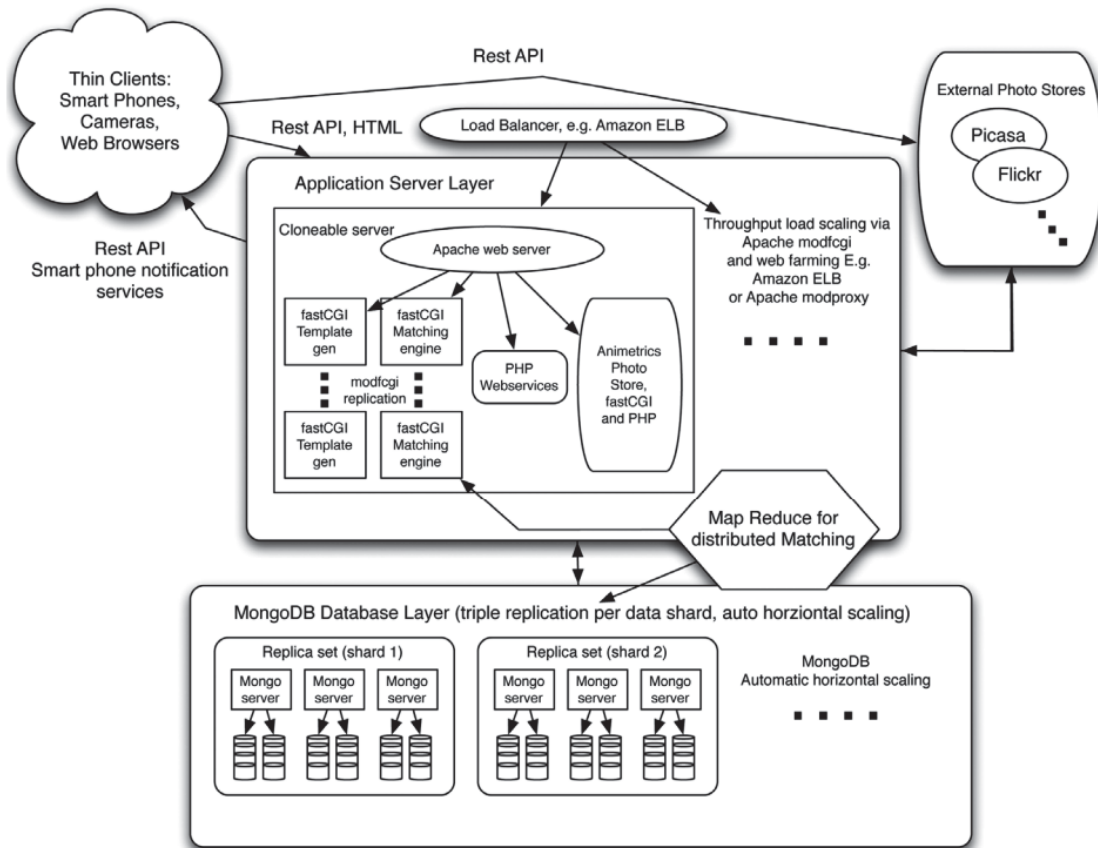
Figure 3: FIMS cloud architecture.

for identification applications (closed set) or watch-list screening (open set). Another relevant application supporting watch-list screening is **FaceR Mobile ID**, which uses iOS and Android mobile devices (over 3G/4G or WiFi) to capture pictures and provide resulting face matches. Besides off-the-shelf applications, Animetrics also provides the **FaceR Facengine SDK** to allow the development of third party applications using FIMS.

Enrolment can be done from still pictures and video frames. Animetrics does not disclose information related to the used features, but their 2D to 3D technology, which the FaceR family is built on, suggests that local physiological features are used (most likely for elastic graph matching). Also, templates are compared as 3D models, to which the vendor attributes improved performance for the recognition task and which allows the recognition of faces from -45 to +45 degrees from just one frontal picture. Memory usage is of 6kb per template and no claims on matching speed are made, but judging from the cloud architecture employed, actual figures may vary depending on the actual server size and hardware. The vendor does not

17

provide information on the video surveillance applications where the technology is used, but operational characteristics suggest that a still-to-still approach is used, extracting faces from video for matching with no actual tracking.

A potential issue with Animetric's approach with FIMS is that, whereas the computational load is limited to the server, network performance may present itself as a potential bottleneck for larger installations requiring real time response, especially when using mobile devices running over WiFi or 3G. Suitable server load balancing should take place to provide better performance. One requirement is that image/face proportion should be at least 1:8 and that at least 64 pixels between eyes are required for optimal performance. This suggests that the product works well with subjects close to the camera, which excludes operation in many video surveillance setups.The vendor does not provide information on product performance or tests performed.

2.6   PittPatt SDK

- Vendor: Pittsburgh Pattern Recognition acquired by Google, Inc.
- Web: http://pittpatt.com
- Type: SDK
- Tasks: face annotation, face identification, enrolment from video to track an individual.

The **PittPatt SDK** was developed by *Pittsburgh Pattern Recognition*, which was recently acquired by *Google Inc.*. The lack of updated information at PittPatt's official website suggests that legacy users of the SDK are still licensed to develop and deploy applications with the SDK, but Google has made no formal announcements as to the future of this technology and downloads are unavailable.

Documentation on the website provides some details on PittPatts inner works. Detection finds faces that are recognized by a frontal face or multi-pose face matcher. Different features are used for each classifier, thus face images on different poses are required for best performance. The face detector also estimates the head pose (roll and yaw). Enrolment can be done from faces detected on video or still images. Template sizes range from 10 kb (frontal faces template) to 120 kb (multiple poses template). Accepted head yaw values are between +18 and -18 degrees for frontal faces, and between +36 and -36 degrees for multiple pose face templates.

Demonstration videos for PittPatt applications are still available online[8], but the vendor makes no performance claims on any available data set. PittPatt participated in a number of still-to-still face recognition NIST-conducted evaluation (see Figure 2), and was the only technology that participated in the NIST video-to-video face evaluation [9].

Whereas the technology was maturing and there were no formal announcements, the acquisition by Google may indicate that the technology will be exclusively used inside Google's services and social networks. Based on these assumptions, using PittPatt with previously issued licenses is not recommended, as

---

[8]http://youtu.be/z76GpB3W-68 and http://youtu.be/wzzuojueKRQ

[9]http://biometrics.nist.gov/cs_links/ibpc2010/pdfs/Phillips_Jonathon_MBE2010-MBGC%20summary%20Mar10.pdf, http://www.biometrics.org/bc2013/presentations/nist_phillips_wednesday_1400.pdf

these versions will become obsolete over time, as updates and fixes are unlikely to happen if Google uses PittPatt as an exclusive technology.

2.7 Genex Technologies SureMatch 3D Suite

- Vendor: Genex Technologies (Technest)
- Web: http://genextech.com
- Type: application, may be customized under request
- Tasks: face identification.

The **SureMatch 3D** technology aims at providing face recognition using 3D models to artificially generate several head poses, an approach similar to the one used by *Animetrics* with **FaceR**. Faces can be captured using either the proprietary **Rainbow 3D** camera for true 3D verification, or traditional 2D cameras. Face enrolment may be done from 2D images, as indicated in Fig. 4, which are mapped to 3D models and allows the conversion of existing watch lists to a 3D format or to an expanded 2D set.

Applications listed by the vendor suggest that the technology is designed for access point control, rather than for video surveillance. Genex provides several off-the-shelf solutions for various scenarios, such as passport check station, major events security checkpoints, etc. Whereas no SDK is provided for custom development, specific customizations may be ordered by clients. No face tracking or tagging capabilities are mentioned, which indicates that this product may not suitable for watch-list screening. The vendor makes generic performance claims, but provide no performance evaluations and reference data-sets to support claims.

SureMatch 3D Recognition Process

| **2D Photo Database** | **Request 2D Photo** | **Convert to 3D** | **Generate Multiple 2D** | **Match 2D Photos** |
| Existing watch list database supplies mug shot | SureMatch 3D queries DB for photograph | 3D FaceMap calculates 3D model | 3D Enroll creates dozens of varying 2D photos | SureMatch 3D sends new watchlist data to 2D database |

Integrated SureMatch 3D Software

Figure 4: SureMatch 3D technology overview.

2.8 FACE-TEK Notiface II

- Vendor: FACE-TEK
- Web: http://face-tek.com
- Type: application

- Tasks: face identification

FACE-TEK's **NotifaceII** is an application that provides limited operational conditions for watch-list screening. It only supports legacy analog CCTV cameras, which provide inferior image quality than current IP surveillance cameras. It offers no tracking support, instead, each time a face is recognized within the watch list, the system displays the face on a separate window to warn the operator and logs the occurrence. The company's profile is dedicated to access control and verification, and it seems that **NotifaceII** is in attempt to capitalize on the analog CCTV legacy market still used in older installations or in under-developed countries.

2.9 Acsys FRS SDK

- Vendor: Acsys Biometrics
- Web: http://www.acsysbiometrics.com
- Type: SDK
- Tasks: face annotation, face identification and enrolment from video to track an individual.

A Canadian company based in Burlington, ON, *Acsys Biometrics* provides both ready-to-use solutions and an SDK to develop face recognition applications, the **Acsys FRS SDK**. The Acsys architecture (Fig. 5) targets several computers inter-connected through a network, all using a central server to obtain and update biometric data, which is replicated asynchronously between clients. The SDK is available for Windows, supporting development on Visual C++ 6.0 (or higher), Visual Basic 5.0 (or higher) and Borland Delphi 5.0 (or higher). The vendor claims that the SDK is also BioApi[10] compliant, complying to certain norms to provide software component modules in order to facilitate integration with other components and hardware. Acsys does not provide solutions for other biometric modalities, thus multi-modal recognition requires the licensing or development of other software components.

The **Acsys FRS SDK** supports enrolment from both still pictures and video sequences. Still picture templates use 4 kb of memory, and video templates use 8 kb. Features extraction relies on the eyes positions, suggesting that local features are used for graph matching. Matching requires a minimum distance between eyes of at least 40 pixels for still images templates and 15 pixels for video templates. Matching speed depends also on the template type, using still image templates allows the matching of 100000 templates per CPU core, whereas using video-based templates allows the matching of 25000 templates per CPU core. Performance claims by Acsys are that the system can perform real time (30 fps) face tracking of 16 simultaneous individuals, but it should be noted that the tracking is performed by a client computer and matching is performed at a centralized server. The Acsys architecture also supports up to 4000 clients over the network and up to a million enrolled users, which by far surpasses specific needs for watch-list screening. Each client can process up to 16 different video feeds, though real-time performance may limit this number to a much lower value.
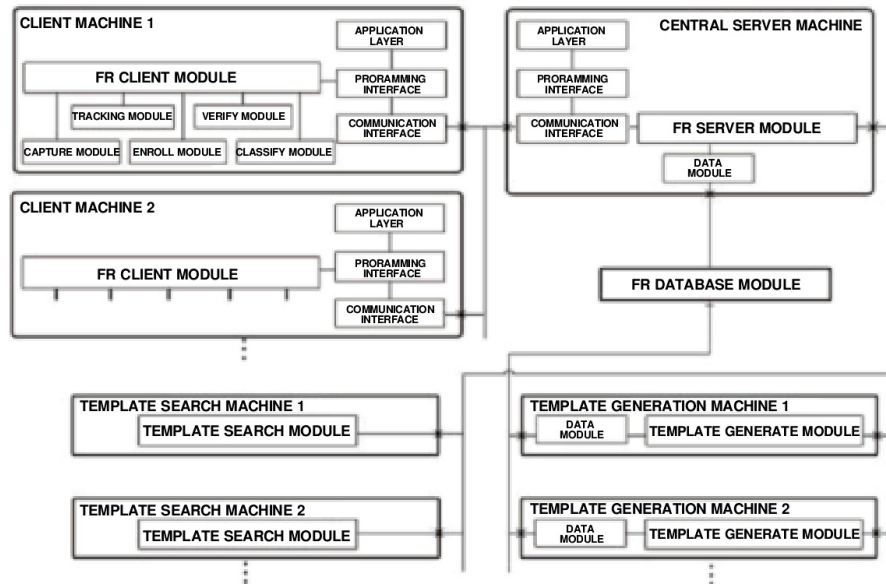
---

[10]http://www.bioapi.org/

Figure 5: Acsys face recognition architecture.

2.10 Airborne Biometrics FaceFirst

- Vendor: Airborne Biometrics Group, Inc.
- Web: http://facefirst.com
- Type: application
- Tasks: face identification.

Developed by *Airborne Biometrics*, **FaceFirst** is a face recognition application built over Cognitec's **FaceVAKS**. It supports enrolment from pictures to build a watch list that is used for screening with IP cameras using the architecture in Fig. 6. A cloud based approach is used to shift the heavy processing to a centralized server. When detected faces are identified at the cloud server, alerts are sent for appropriate action on different devices, like desktop surveillance stations or smartphones through SMMS, including the option to filter alerts to specific channels.

Cognitec's **FaceVaks** allows face annotation and enrolment from live video for tracking, but FaceFirst's description makes no mentioning of those features. The lack on the website of more detailed information suggests that **FaceFirst** is a product deployed according to specific customizations required for each client installation. Performance figures should be consistent with those provided by Cognitec's FaveVACS.
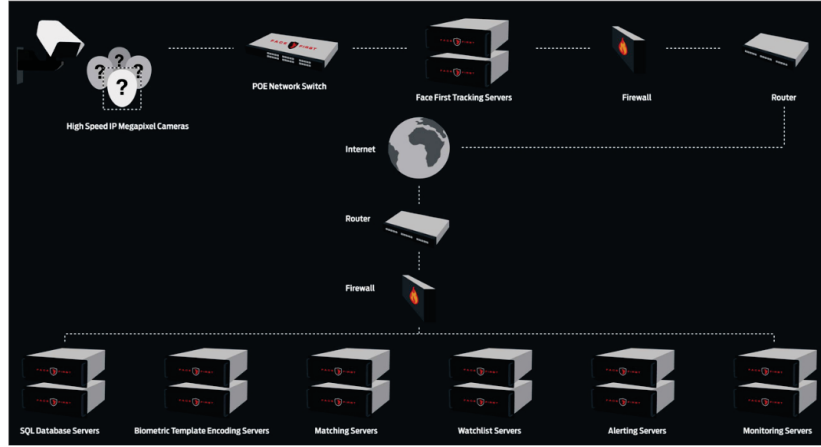
Figure 6: FaceFirst architecture.

3 Patents related to face recognition in video

Table 3 summarizes patents relevant to FR in video. In addition to brief description, each patent is presented with its filing date, which allows us to determine the patent's expiration date. Unlike commercial products, patents have no associated performance metrics as they target the description of a method. Patents that may potentially conflict with possible deployment of FR in video surveillance applications receive particular attention.

Table 3: Patents related to face recognition in video.

Patent	Agency	Number	Filing	Issued	Assignee
Autonomous Face Recognition	United States Patent	5,012,522	1988/12/08	1991/04/30	US Air Force
Method and apparatus for image recognition using invariant feature signals	United States Patent	5,497,430	1994/11/07	1996/03/05	Physical Optics Corporation (Torrence, California)
Face annotation in streaming video	United States Patent	US2008/0235724 A1	2008/03/25	2006/09/16	Koninklijke Philips Electronics, N.V.
Method and System for Automated Annotation of Persons in Video Content	United States Patent	US2010/0008547 A1	2010/01/14	2008/07/15	Google Inc.
Open set recognition using transduction	United States Patent	7,492,943 B2	2009/02/17	2005/03/10	George Mason Intellectual Properties Inc.
Open set recognition using transduction	United States Patent	7,492,943 B2	2009/02/17	2005/03/10	George Mason Intellectual Properties Inc.
Method for incorporating facial recognition technology in a multimedia surveillance system	United States Patent	US 7,634,662 B2	2009/12/15	2003/11/21	David A. Monroe
Combined face and iris recognition system	United States Patent	US 2008/0075334 A1	2008/03/27	2007/03/02	Honeywell International Inc.
Method for Robust Human Face Tracking in Presence of Multiple Persons	United States Patent	US 6,404,900 B1	2002/06/01	1999/01/14	Sharp Laboratories of America Inc.
Face Recognition from Video Images	Canadian Intellectual Property Office	CA 2326816	2005/04/05	1999/04/12	Google Inc.
Face Identification Apparatus and Entrance and Exit Management Apparatus	Canadian Intellectual Property Office	CA 2537738	–	2006/06/27	KABUSHIKI KAISHA TOSHIBA
Automatic Biometric Identification Based on Face Recognition and Support Vector Machines	United States Patent	US 2009/0074259 A1	–	2005/07/29	Inventors
Facial Recognition System and Method	United States Patent	US 7,643,671 B2	2010/01/05	2004/01/21	Animetrics Inc.

3.1 Autonomous Face Recognition

- Agency: United States Patent
- Number: 5,012,522
- Filing: 1988/12/08
- Issued: 1991/04/30
- Assignee: US Air Force
- Inventor: Laurence C. Lambert

This patent introduces a machine that can autonomously locate and recognize faces in video scenes with random content within two minutes. The machine uses images obtained from a video camera and is insensitive to variations in brightness, scale, focus and require no human intervention or input. A suggested embodiment of this system uses a camera, a Micro-Vax computer, an A/D converter and a hardware printout. The computer role (Fig. 7) is to perform a pattern recognition algorithm to search facial components, identify a gestalt face [11] and compare it to a stored set of facial characteristics of known human faces.

3.2 Recognition System – Particularly for Recognizing People

- Agency: United States Patent
- Number: 5,412,738
- Filing: 1993/08/10
- Issued: 1995/05/02
- Assignee: Instituto Trentino di Cultura (Trento, Ialy)
- Inventors: Roberto Brunelli, Daniele Falavigna, Tomaso Poggio, Luigi Stringa

This patent describes a multi-modal biometric system (Fig. 8), which recognizes or identifies persons using acoustic and visual features. The system described in this patent is not related to unconstrained surveillance tasks, but is similar to the scenario of individual surveillance at the primary inspection line (PIL) in airports.

3.3 Face Annotation in Streaming Video

- Agency: United States Patent
- Number: US2008/0235724 A1
- Filing: 2006/09/16

---

[11]Gestalt psychologists theorize that a face is not merely a set of facial features but is rather something meaningful in its form. This is consistent with the Gestalt theory that an image is seen in its entirety, not by its individual parts. Hence, the "gestalt face" refers to a holistic representation of face. Gestalt's theory "Figure and Ground" defines relevant characteristics of a face from pictures. http://en.wikipedia.org/wiki/Face
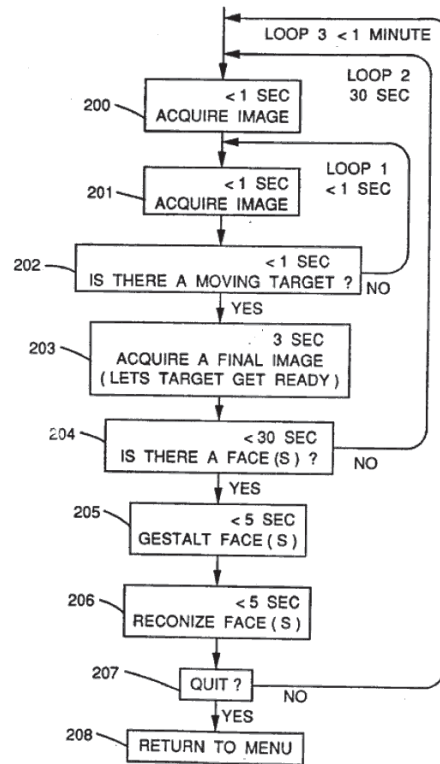
Figure 7: Autonomous face recognition algorithm.

- Issued: 2008/03/25
- Assignee: Koninklijke Philips Electronics, N.V.
- Inventors: Frank Sassenscheidt, Christian Nenien, Reinhard Kneser

This patent describes a system and a method for detecting and annotating faces on-the-fly in video data. Annotation is performed by modifying pixel content and is independent of file types, protocols or standards. The invention can also perform real-time face recognition, comparing detected faces to known identities to add personal information to annotations. The invention is described for applications related to video-conferences on classrooms and meetings, but has enough similarities to face annotation application in video surveillance, as one IP surveillance camera streams video over a network for video processing on a remote computer.

3.4 Method and System for Automated Annotation of Persons in Video Content
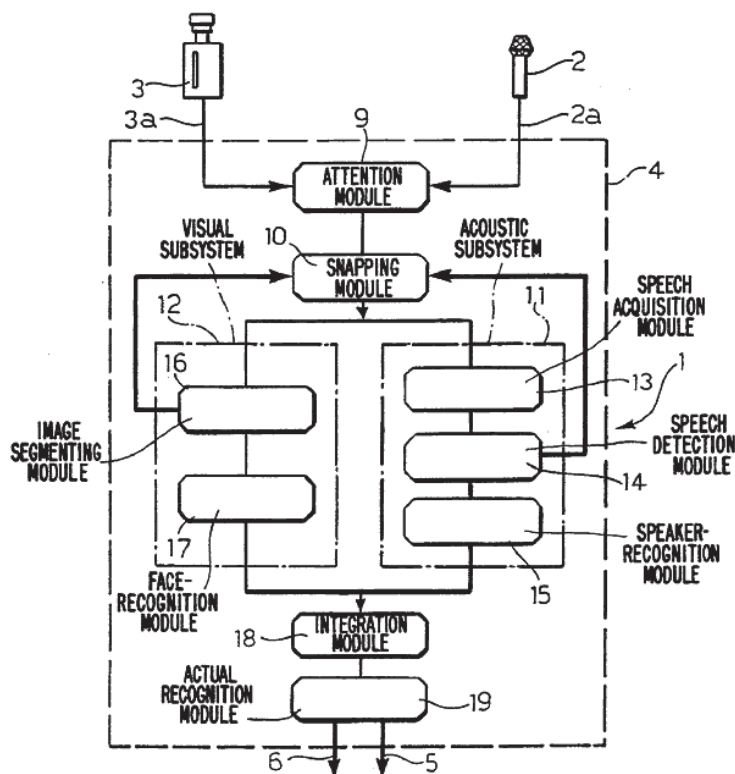
- Agency: United States Patent

Figure 8: Overview of the multi-modal biometric recognition system.

- Number: US2010/0008547 A1
- Filing: 2008/07/14
- Issued: 2010/01/24
- Assignee: Google Inc.
- Inventors: Jay Yagnik, Ming Zhao

Google Inc. filled this patent to automatically annotate online video content, an effort directed to the company's video services on YouTube and on the Google+ social network. It is most likely related to the recent acquisition of Pittsburgh Pattern Recognition (PittPat). The patent describes one embodiment where a computer-implemented method that identifies faces in a video, generating face tracks that are clustered, where each face cluster is associated to one or more key face images that are correlated to faces stored in a faces database. This application scenario is exactly the same as demonstrated on PittPat's demonstration videos. This patent hints to the application of the PittPat SDK by Google Inc. and is relevant to the video

surveillance in a scenario where an operator needs to analyze video surveillance footage for the presence and motion patterns of persons of interests.

3.5 Open Set Recognition using Transduction

- Agency: United States Patent
- Number: 7,492,943 B2
- Filing: 2005/03/10
- Issued: 2009/02/17
- Assignee: George Mason intellectual Properties Inc.
- Inventors: Fayin Li, Harry Wechsler

The open set Transduction Confidence Machine kNN (TCM-kNN) algorithm [5] is the focus of this patent, which presents a classification method for open-set problems, such as in automated watch-list video screening application. Any application that uses the TCM-kNN algorithm has a high risk of infringing on this patent. More information on the TCM-kNN algorithm is provided in report [1].

3.6 Method for incorporating facial recognition technology in a multimedia surveillance system

- Agency: United States Patent
- Number: US 7,634,662 B2
- Filing: 2003/11/21
- Issued: 2009/12/15
- Assignee: David A. Monroe, 7800 IH-10 West, #700, San Antonio, TX (US) 78230.
- Inventors: David A. Monroe

The basic embodiment of this patent is detailed in Fig. 9, where a camera connected to an IP network views a scene of interest, and a processor analyzes the video and performs facial separation and face signature generation (feature extraction) for database look up in order to trigger alarms when appropriate. This scenario is very similar to the applications used in video surveillance, but is also very generic and of broad scope.

3.7 Combined face and iris recognition system

- Agency: United States Patent
- Number: US 2008/0075334 A1
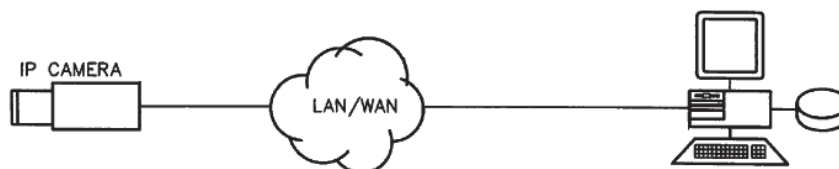- Filing: 2007/03/02
- Issued: 2008/03/27

Figure 9: Face recognition in a multimedia surveillance system.

- Assignee: Honeywell International Inc., Morriston, NJ
- Inventors: Gary E. Determan, Vincent C. Jacobson, Jan Jelinek, Thomas Phinney, Rida M. Hamza, Terry Ahrens, George A. Kilgoe, Rand P. Whillock, Saad Bedros

The invention describes a bi-modal biometric recognition system, based on video cameras that can capture both the users face and iris images at a distance, providing the visual interface shown in Fig. 10, where target individuals have both their face and iris used to confirm their identity with higher confidence. However, technical limitations indicate that the invention targets future technological advancements, since capturing iris images at a distance is difficult with currently existing technologies.
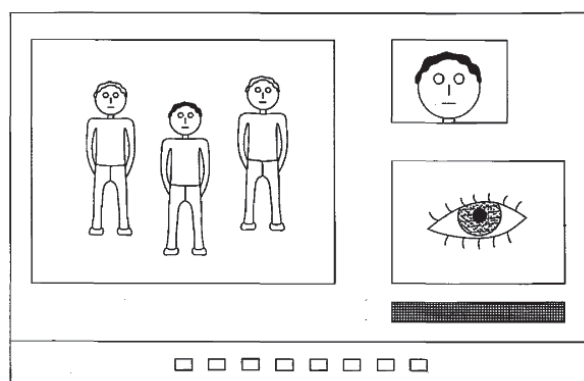

Figure 10: Combined face and iris recognition system user interface.

3.8 Method for Robust Human Face Tracking in Presence of Multiple Persons

- Agency: United States Patent
- Number: 6,404,900 B1
- Filing: 1999/01/14
- Issued: 2002/06/11

- Assignee: Sharp Laboratories of America Inc.
- Inventors: Richard J. Qian, Kristine E. Mathews

The invention describes a method of using color-based filtering in combination with a motion estimation technique, using a linear Kalman filter. The invention provides an improved method to track a dominant face and is insensitive to partial occlusions, shadows, face orientation, changes in scale and lighting conditions. Motion of a tracked face is modelled as a constant 2D translation within the image plane to estimate the face position on subsequent frames and verify its occurrence.

3.9 Face Recognition from Video Images

- Agency: Canadian Intellectual Property Office
- Number: CA 2326816
- Filing: 1999/04/12
- Issued: 2005/04/05
- Assignee: Google Inc.
- Inventors: Thomas Maurer, Egor V. Elagin, Luciano P.A. Nocera, Johannes B. Steffens, Harmut Neven

The patent describes a method for detecting and recognizing objects in images. The method uses a two-stage approach for recognition of faces or objects. The first stage detects and tracks objects in video frames using computationally efficient algorithms (without mentioning any specific algorithm), while the second stage uses elastic bunch graph matching to identify the face (or object). The idea behind the two-level approach (according to inventors) is to overcome the limitation posed by the elastic bunch graph matching, which does not perform well when the target object occupies only a small fraction in the image. The approach covers both traditional monocular videos, as well as stereo video sequences.

3.10 Face Identification Apparatus and Entrance and Exit Management Apparatus

- Agency: Canadian Intellectual Property Office
- Number: CA 2537738
- Filing: 2006/06/27
- Issued: not yet issued
- Assignee: KABUSHIKI KAISHA TOSHIBA
- Inventors: Kei Takizawa

The system detailed in Fig. 11 is not related directly to face recognition applications, but to a solution to capture faces in a security checkpoint scenario. The patent proposes a method to efficiently capture faces of a walking person, including covert operation. This checkpoint scenario is similar to the primary inspection

line in international airports, where a person walks towards the checkpoint and is interviewed by the office. The inventor detailed the scenario where the system automatically let the person in or out depending on the identification results, which can be associated to a decision support system at an immigration booth at the primary inspection line.
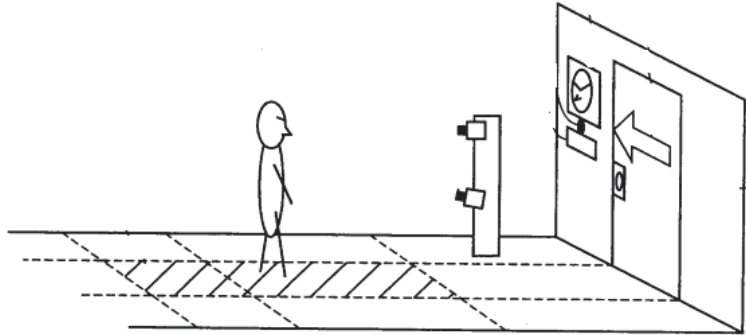

Figure 11: Face identification to control an entrance and exit checkpoint.

3.11 Automatic Biometric Identification Based on Face Recognition and Support Vector Machines

- Agency: United States Patent
- Number: US 2009/0074259 A1
- Filing: 2005/07/29
- Issued: not yet issued
- Assignee: the inventors
- Inventors: Madalina Baltatu, Rosalia D'Alessandro, Roberta D'Amico, Massimo Tistarelli, Enrico Grosso, Manuele Bicego

The patent describes face biometric identification on images, including, but not limited to, live video streams or saved video files. The method uses the support vector machines (SVM) classifier and the enrollment/identification procedure, and consists of capturing multiple biometric samples that go through face detection to extract region of interests that have their features extracted. In the case of enrolment, a supervised procedure is used to tag the user in the system for proper training. During the identification stage, the features are used to obtain scores from the trained classifiers. This method describes the use of the OpenCV library for video capture, the Viola Jones algorithm to detect faces and an Radial Basis Function (RBF) SVM classifier for face identification. All these components are readily available without cost and, if issued, this patent may pose problems to embodied methods in specific systems that use the same components.

3.12 Facial Recognition System and Method

- Agency: United States Patent
- Number: US 7,643,671 B2
- Filing: 2004/01/21
- Issued: 2010/01/05
- Assignee: Animetrics Inc.
- Inventors: Kenneth Dong and Elena Dotsenko

Animetrics uses this patent in their FaceR product, discussed in the previous section. The described method uses 3D facial models to correct the user pose (so it is always facing the camera) and lighting issues like shadows, to improve face recognition. The face recognition stage is generic and not relevant, but the region of interest correction is relevant to video surveillance in unconstrained environments. The approach would allow the correction of Regions of Interest (ROIs) from users not facing the camera and improve recognition rates. However, we assume that the method is computationally intensive. Besides extracting the region of interests and performing recognition itself, the method also requires processing time to estimate the direction the face is looking to in order to map the face to a 3D model, remove shadows and provide the final image for recognition. Thus, the system is feasible to perform recognition on still images, but for real time video stream or crowded scenes it might take some time until cost effective hardware becomes available.
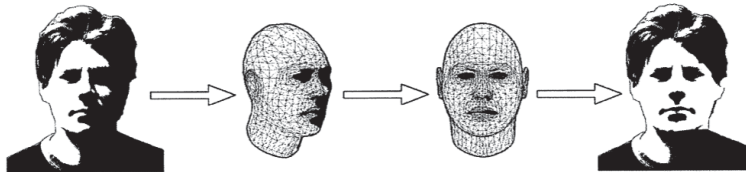

Figure 12: Pose and lighting corrections through 3D models.

4 Discussion

Commercial off-the-shelf FR matching software provide means for in-house development of customized FR solutions for video surveillance applications. To better understand the properties and limitations of FR matching technologies, it is recommended to conduct the evaluation of these technologies using the multi-level methodology proposed in report [3].

Based on the availability and affordability of vendor FR SDKs at the time of conducting this study, the following three COTS products have been selected for evaluation (see also Table 4), with the results being presented in report [4]:

- **VeriLook Surveillance SDK** (http://www.neurotechnology.com) – based on its affordability, its features and also on the availability of complementary SDKs from the same company for multi-modal biometrics.

- **PittPatt SDK** (http://pittpatt.com/) – based on its availability, its reported performance on video data, and capability to enroll and process faces from both still and video data.

- **FaceVACS SDK** (http://cognitec-systems.de), - – based on its availability, high performance on still facial images and capability to enroll and process faces from both still and video data.

Table 4: Commercial products evaluated within the study.

Name	Company	Notes
VeriLook Surveillance SDK	NEUROtechnology http://neurotechnology.com/	Support multi-modal biometrics. 60000 template matches per second.
PittPatt SDK	Google http://pittpatt.com/	Both still-to-video and video-to-video.
FaceVACS SDK FaceVACS-Video Scan API	Cognitec Systems http://cognitec-systems.de	Support multi-modal biometrics. 142000 template matches per second.

In addition to their SDK, these companies also provide ready to use solutions for video surveillance applications, which were not tested within this study, but which can be recommended for evaluation. Specifically, Cognitec provides also the **FaceFirst** surveillance solution, which is ready to use and is built on the **FaceVACS SDK**.

Another product that was not evaluated during this study but which can be recommended for evaluation is **FaceR Engine SDK**. Whereas the architecture offers a distributed environment with load balancing and an impressive customer list (including several government agencies), it is also the most restrictive SDK, as it is geared towards face recognition in static pictures and would not allow the quick development of a video surveillance system. However, for traditional face recognition applications applications, like confirming identities from pictures obtained with mobile devices, the FaceR Engine should be considered.

Finally, NEC NeoFace SDK, which is the best performing in low resolution images according to the latest two NIST Face Recognition Vendor Test results [6], appears to be the best candidate for prototyping FR solutions for video surveillance applications, where faces are typically of low resolution. As highlighted in the introduction however, since face matching is only one of several face processing components that need to be developed for successful solution, using the best performing face matcher does not necessarily yield the best recognition performance. Furthermore, as highlighted in several other work [1, 3], face recognition solutions for video surveillance applications can be significantly improved by using video analytics and face tracking, implying that less performing face matching products can still be used for developing the prototypes.

With respect to the patent survey, we have analyzed face recognition patents for their relevance to video surveillance applications. The patent which was found the most restrictive relates to the TCM-kNN algorithm (US patent 7,492,943 B2), as it patents the algorithmic procedure and requires a license to use the classifier. Other patents do not show high relevance to the video surveillance applications studied in this project. Nevertherless, they may still need to be taken into account when developing a customized face recognition solution for video surveillance applications to reduce the risk of intellectual property infringement.

References

[1] E. Granger, P.Radtke, and D. Gorodnichy, "Survey of academic research and prototypes for face recognition in video," CBSA, Tech. Rep., 2014.

[2] G. W. Q. Patrick J. Grother and P. J. Phillips, "Report on the evaluation of 2d still-image face recognition algorithms," NIST Interagency Report, Tech. Rep. 7709, 2010.

[3] E. Granger and D. Gorodnichy, "Evaluation methodology for face recognition technology in video surveillance applications," CBSA, Tech. Rep., 2014.

[4] D. Gorodnichy, E. Granger, E. Choy, W. Khreich, P.Radtke, J. Bergeron, and D. Bissessar, "Results from evaluation of three commercial off-the-shelf face recognition systems on chokepoint dataset," CBSA, Tech. Rep., 2014.

[5] F. Li and H. Wechsler, "Open set fr using transduction," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 11, pp. 1686 –1697, nov. 2005.

[6] P. G. M. Ngan, "Face recognition vendor test (frvt). performance of face identification algorithms," NIST Interagency Report, Tech. Rep. 8009, 2014.